Bayesian Hierarchical Models to Augment the Mediterranean Forecast System

Ralph F. Milliff
Colorado Research Associates Division
NorthWest Research Associates Inc.
3380 Mitchell Lane, Boulder, CO, USA 80301-5410
phone:303-415-9701 x216 fax:303-415-9702 email: milliff@cora.nwra.com

Mark Berliner
Department of Statistics
Ohio State University
1958 Neil Avenue, Columbus, OH, USA 43210-1247
phone:614-292-0291 fax:614-292-2096 email: mb@stat.ohio-state.edu

Christopher K. Wikle
Department of Statistics
University of Missouri-Columbia
146 Middlebush, Columbia, MO, USA 65211
phone:573-882-9659 fax:573-884-5524 email: wikle@stat.missouri.edu

Award Number: N00014-05-C-0198
http://www.cora.nwra.com/~morzel/MED.BHM/med.bhm.html

LONG-TERM GOALS

Eighteen months into the project, the long-term goals and objectives remain as stated in the progress report last year. Our goal is to develop an ensemble ocean forecast methodology, using Bayesian Hierarchical Modelling (BHM) tools. The ocean ensemble forecast methods to be developed should be practical enough to benefit the Mediterranean Forecast System (MFS) in its operational mode, and they should demonstrate forecast uncertainties during difficult to predict regime transitions in the Mediterranean Sea (e.g. the Fall transition, deep water formation).

OBJECTIVES

Two main objectives comprise the research plan.

First, an ensemble of ocean initial conditions will be derived from realizations of the surface wind forcing as drawn from a posterior distribution of a BHM for the surface wind process. The surface wind field realizations have been used in separate data assimilation steps to produce unique, but realizable ocean initial conditions. The surface wind likelihood distributions are based on QuikSCAT data and ECMWF analyses. The prior distributions are based on a time-dependent augmentation of the stochastic geostrophy model introduced by Royle et al. (1998).

The second objective involves the accurate representation of forecast error covariance evolution in MFS.

The operational implementation of a reduced order optimal interpolation (ROOI) data assimilation method for MFS involves an ad-hoc truncation in the representation of the background error

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **30 SEP 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Bayesian Hierarchical Models to Augment the Mediterranean Forecast System**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **NorthWest Research Associates Inc,Colorado Research Associates Division,3380 Mitchell Lane,Boulder,CO,80301-5410**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

covariance. BHM can be used to remove this arbitrary truncation. Moreover, as ensemble forecasts are run in MFS for abrupt seasonal transition events, the ensemble spread will be used to refine priors in the BHM for error covariance evolution.

APPROACH

The BHM theory and fundamentals of BHM implementation for the ensemble forecast problem were described in the ONR progress report for FY2005. Here we describe specifics of our approach with respect to a) the ensemble initial condition generation based on surface wind field realizations from a surface wind BHM (MFS-Wind-BHM); and b) BHM to evolve background error covariance in the MFS ROOI (MFS-Error-BHM).

a) MFS-Wind-BHM

Figure 1 shows 10 SVW field realizations drawn from the posterior distribution for MFS-Wind-BHM at a single time during the data assimilation phase of a reforecast experiment run for the period 1 through 24 February 2005. This period was chosen because it corresponds to a strong Mistral event, and deep-water formation (DWF) response, in the Gulf of Lions. The blue SVW clusters provide a pictorial representation of SVW ensemble variance, or "spread" at each MFS-Wind-BHM output grid location in a blow-up of the western Mediterranean Sea, centered on the Gulf of Lions (Fig 1).

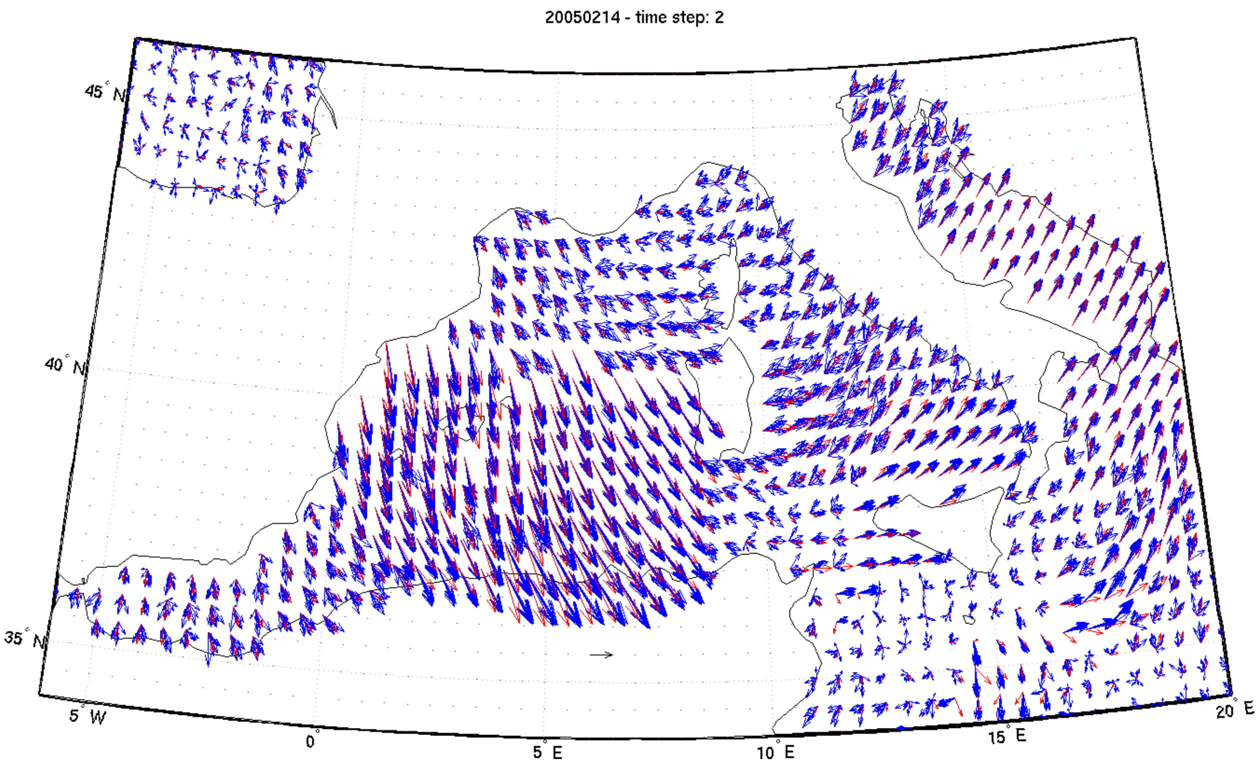

Figure 1. Surface vector wind field realizations for the western Mediterranean basin. Ten blue wind vectors are plotted at each MFS-Wind-BHM output grid location for this snapshot. A red vector at each location represents the surface wind analysis from ECMWF for this time. The reference vector at about 6ºE, 36ºN is scaled to 10 ms-1. The snapshot corresponds to the penultimate ensemble forcing field during the data assimilation period for the reforecast experiment (day 14A). A low pressure system is deepening over Sardinia, and the Mistral winds are freshening to the west and southwest. Where wind speeds are large, the ensemble winds are more narrowly distributed about the ensemble mean wind. In sheltered regions (i.e. near the Sicily Strait), the wind speeds are low and direction uncertainty is larger. The ten realizations at each grid location are all consistent with the uncertainties expressed in the data stage distribution, based on ECMWF analysis and scatterometer observations, and with the process model stage distribution.

Separate data assimilation steps (14 days) and subsequent forecasts (10 days) were performed for each SVW field realization drawn from the MFS-Wind-BHM posterior distribution.

b) MFS-Error-BHM

The ROOI data assimilation system for MFS (called SOFA; System for OceanForecast and Analysis) is described in De Mey and Benkiran (2002), Demirov et al. (2003) and Dobricic et al. (2006). In the optimal interpolation approximation to the Kalman Filter, a time-step by time-step evolution of the error covariance process is replaced by a background error covariance matrix **B**. The MFS implementation partitions the Mediterranean into 13 regions for which a single regional covariance average **B** is used (Dobricic et al., 2006). The vertical structure of the MFS **B** is further approximated in a decomposition onto the set of 20 leading EOFs of the multivariate vertical error covariance matrix:

$$\mathbf{B} = \mathbf{Z}\,\mathbf{B_r}\,\mathbf{Z}^T$$

$$= \mathbf{Z}\,\mathbf{\Lambda}^{1/2}\,\mathbf{C}\,\mathbf{\Lambda}^{1/2}\,\mathbf{Z}^T$$

Here, the multivariate vertical EOFs are in **Z**, and $\mathbf{B_r}$ contains horizontal covariances and eigenvalues for the EOFs, as noted in the second equation.

In MFS-Error-BHM we seek to a) improve vertical covariances in the error process and b) develop a methodology for more highly resolved temporal evolution of **B**. We define a state vector comprised of temperature and salinity vertical error profiles for each region. Two forms of data will comprise our data stage distributions. In one, we use the observation/forecast misfits obtained principally from ARGO floats. Let these misfits be given by $\mathbf{d} = \mathbf{H}\mathbf{x}^f - \mathbf{x}^{ARGO}$, where **H** is the observation operator that moves the model forecast $\mathbf{x}^f$ to the ARGO location for each observation time. A second data set will be constructed by comparing the forecast with the long-term average forecast for the same day. Let $\mathbf{q} = \mathbf{x}^f - \mathbf{x}^{avg}$. Then the data stage distributions for MFS-Error-BHM can be written:

$$\mathbf{d}_t \mid \mathbf{e}_t \sim N(\mathbf{H}_{\mathbf{d}t}\,\mathbf{e}_t, \sigma_d^2\,\mathbf{I})$$

$$\mathbf{q}_t \mid \mathbf{e}_t \sim N(\mathbf{H}_{\mathbf{q}t}\,\mathbf{e}_t, \sigma_q^2\,\mathbf{I})$$

where **e** is the true error process in the vertical, and the σ^2 are *variances* about the error process to be specified at higher levels of the MFS-Error-BHM hierarchy.

We assume we know a robust vertical basis set **U** to decompose the vertical error process at any time, within any of the13 regions of the MFS domain. Given this,two levels of a candidate process model distribution for MFS-Error-BHM can be written:

$$\mathbf{e}_t \mid \mathbf{B_r}_t \sim N(0, \mathbf{U}\,\mathbf{B_r}_t\,\mathbf{U}^T)$$

$$\mathbf{B_r}_t \sim W^{-1}((\upsilon B), \upsilon)$$

where $W^{-1}()$ denotes an Inverse-Wishart distribution (i.e. an inverse gamma equivalent for multi-variate normal distributions (e.g. see Evans et al., 2000). Note that this is a model for *time-dependent* $\mathbf{B_r}_t$. In what might be an easier alternative implementation, we will explore the possibility of isolating time-dependence in the diagonal matrix $\mathbf{\Lambda}$ (from above); i.e. we assume $\mathbf{U_t} \sim \mathbf{Z}\,\mathbf{\Lambda}_t^{1/2}$.

WORK COMPLETED

The prototype MFS-Wind-BHM was designed and implemented based on *stochastic geostrophy* concepts first described by Royle et al., (1998)

Milliff visited INGV Bologna for 30 days in October 2005 . The purpose of the visit was to establish reforecast experiment procedures, scope the computing resources necessary to carry out reforecast experiments, and to identify INGV personnel who would be involved with MFS-Wind-BHM.
Pilot reforecast experiments in MFS-Wind-BHM were run while Milliff was in Bologna.

Mr. Alessandro Bonazzi, a PhD candidate Univ. Bologna, worked with Milliff at INGV and visited NWRA/CoRA in Boulder for 6 months; March-August 2005. Mr. Bonazzi learned principles and

mechanics of BHM, and performed day-to-day tasks in MFS-Wind-BHM reforecast experiments. Mr. Bonazzi is responsible for the coding refinements we implemented on the prototype surface wind BHM.

A DWF event in February 2005, in the Gulf of Lions, was identified for reforecast experiments. A large suite of reforecast experiments has been run using 14-day analyses and 10-day reforecast periods around the DWF event.

Preliminary results of the MFS-Wind-BHM deep water formation experiments were presented in a seminar at ONR headquarters in May 2006. Prof. Pinardi (U. Bologna/INGV) traveled to Washington to co-deliver this seminar with Milliff.

The second "All Hands" meeting for the project was held at NWRA/CoRA in July 2006. Details of the MFS-Error-BHM were scoped, collaborating INGV personnel were identified, and we learned of generalized methods for ocean forecast diagnostics (from Prof. A. Moore, UCSC) as a means of characterizing our reforecast experiment results.

RESULTS

Figure 2 demonstrates the spread in ocean response in terms of the standard deviations (computed over 10 ensemble members) of sea surface height (SSH; top panel) and sea surface temperature (SST; bottom panel). The variance in both fields concentrates in ocean mesoscale eddy structures that are the uncertain elements of the ocean forecast system. The localization of ocean response uncertainty in the mesoscale persists through the forecast period (10 days) as well. Figure 3 depicts the ensemble mean SSH forecast (top panel) and SSH standard deviation of the ensemble forecast (bottom panel) on day 10 of the forecast.

The concentration of initial condition and forecast ensemble variance in the ocean mesoscale is a siginificant achievement for the MFS-Wind-BHM methodology. The ensemble spread in SVW implies ensemble variances in wind stress curl and Ekman pumping at the surface. The variance in surface vorticity forcing appears to drive directly a variable mesoscale eddy response in the upper ocean.

The mesoscale coherent structures in the Mediterranean occur on spatial scales of tens of kilometers at the surface (Figs 2 and 3). These features are not temporally or spatially resolved within an assimilation cycle. Because they interact with the sub-basin scale circulations that combine to form the Mediterranean general circulation (Fig. 3, top panel), the mesoscale eddy field comprises the largest uncertain component of the MFS analyses and forecasts. The MFS-Wind-BHM initialization builds variance in the mesoscale in physically consistent ways, using an ensemble of realizable surface vector winds.

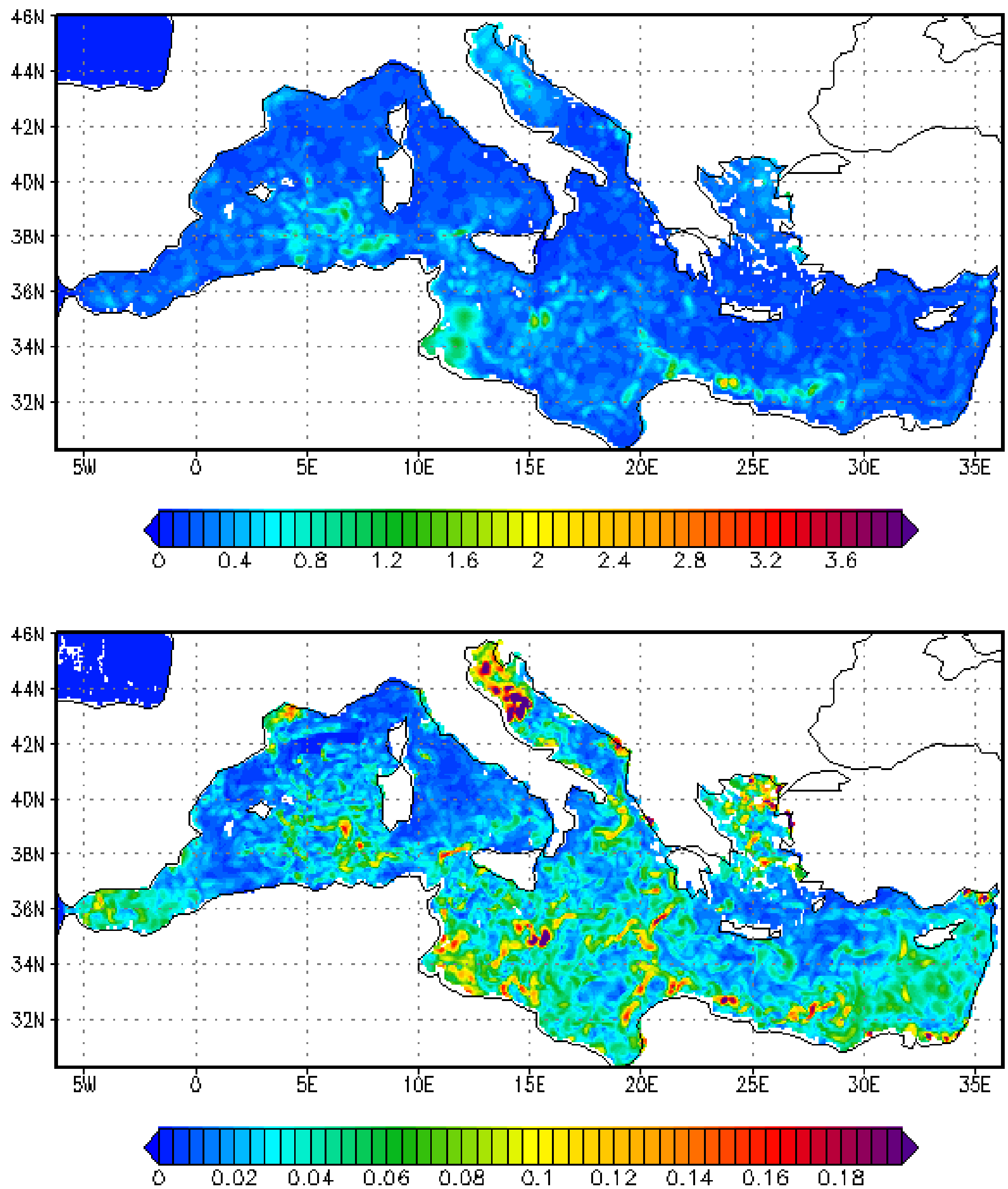

Figure 2. Ensemble forecast initial condition spread for SSH (top panel; cm) and SST (bottom panel; ° C). The standard deviation of the SSH and SST are computed at day 14 of the data assimilation period (day 0 of the forecast) across the 10 fields that have been driven during the data assimilation by realizations from MFS-Wind-BHM (Fig. 1). Ensemble spread is concentrated in ocean mesoscale eddies. SSH initial condition uncertainty amplitudes in the mesoscale (top panel) reach about 1.5 cm, with maxima located in the Algerian Current region south of the Gulf of Lions, and off the North African Coast in the regions of the Tunisian Plateau and the Herodotus Trough. SST initial condition uncertainty amplitudes (bottom panel) range as high as 2° C in the Northern Adriatic, and in isolated eddies in the Ionian Sea. SST uncertainty in mesoscale structures is much more widely spread than is the case for SSH uncertainty. Relative minima in SST initial condition uncertainty occur east of Majorca, in the northern Tyrrhenian Sea, and in the region of the Rhodes Gyre.

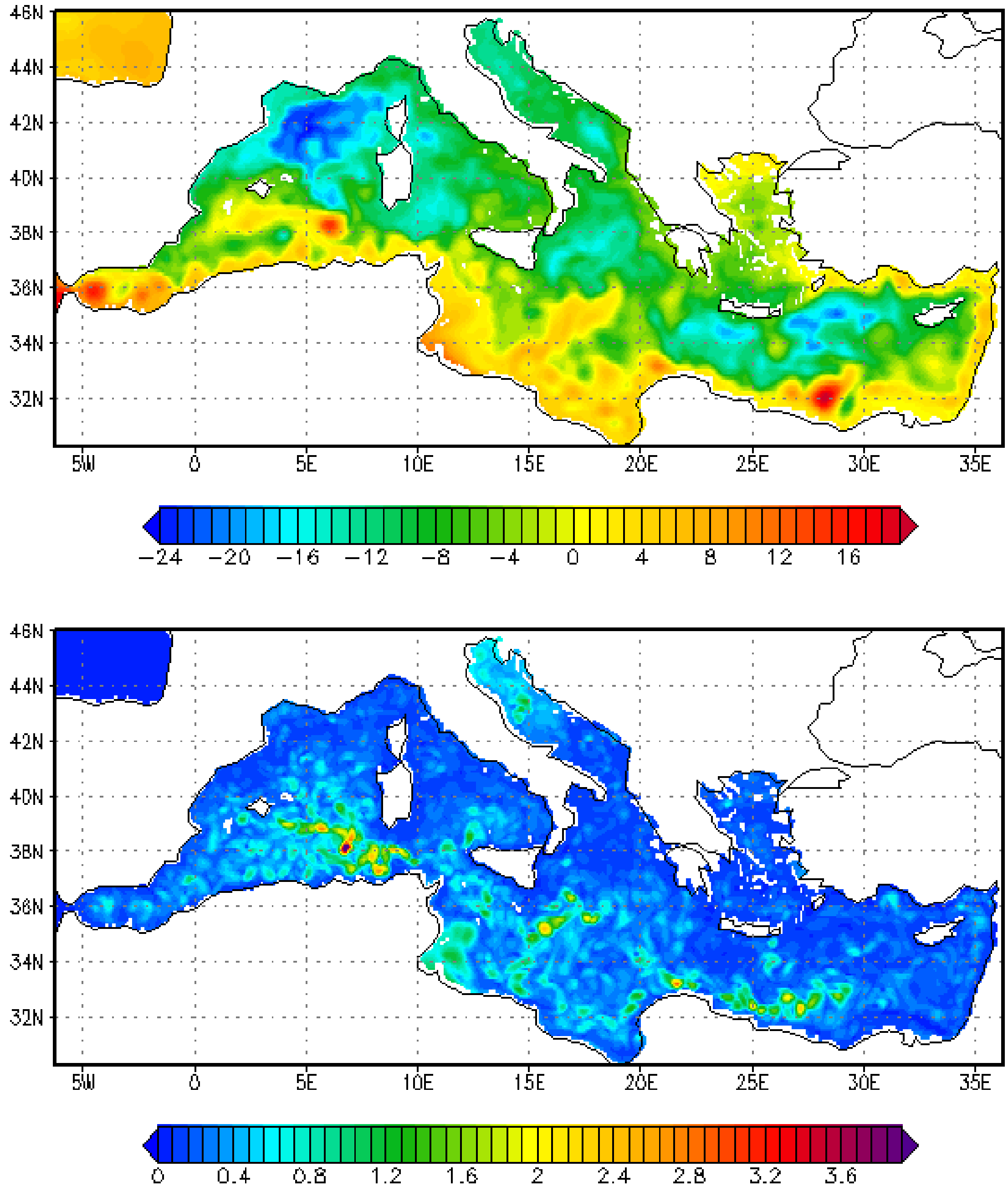

Figure 3. Ensemble forecast ensemble mean SSH (top panel; cm) and ensemble standard deviation (bottom panel; cm) from day 10 of the forecast. Ten forecasts were driven by surface wind realizations from MFS-Wind-BHM. Ensemble spread (standard deviation) is concentrated in ocean mesoscale eddies, and increases in amplitude throughout the 10d forecast period. The general circulation features evident in the ensemble mean SSH include: boundary currents along the North African coast (e.g. the Algerian Current); the Gulf of Lion Gyre (preconditioned for DWF); the Ionian-Atlantic current; the Mersa Matruh Gyre; and evidence of the Rhodes Gyre and related cyclones in the Northern Levantine Sea. The SSH ensemble spread (bottom panel) reaches amplitudes of more than 3 cm in mesoscale eddies associated with instabilities of the Alegerian Current. These eddies impact the DWF event in the Gulf of Lions to the North.

RELATED PROJECTS

“Bayesian Hierarchical Models to Augment the Mediterranean Forecast System” (same title), Prof. L. Mark Berliner (Principal Investigator), Department of Statistics, Ohio State University, Grant Number N00014-05-1-0336; Prof. Christopher K. Wikle (Principal Investigator), Department of Statistics, University of Missouri, Grant Number N00014-05-1-0337.

“The Mediterranean Forecast System: Toward Environmental Prediction”, Prof. Nadia Pinardi, Scientific Coordinator, Istituto Nazionale di Geofisica e Vulcanologia (National Institute of Geophysics and Volcanology) , EU Contract Number EVK3-CT-2002-00075

“Continued Development of 4D-Variational Data Assimilation and Adjoint-Based Methods of Sensitivity Analysis and Applications Using ROMS”, Andrew M. Moore (Principal Investigator), Department of Ocean Sciences, University of California, Santa Cruz, Grant Number N00014-06-1-0406.

REFERENCES

De Mey, P. and M. Benkiran, 2002: “A multi-variate reduced order optimal interpolation method and its application to the Mediterranean basin-scale circulation” in *Ocean Forecasting: Conceptual basis and applications*, N. Pinardi and J. Woods (eds.), Springer-Verlag, 281-305.

Demirov, E., N. Pinardi, C. Fratianni, M. Tonani, L. Giacomelli, and P. De Mey, 2003: “Assimilation scheme of the Mediterranean Forecasting System: Operational Implementation.”, *Annales Geophysicae*, **21**, 189-204.

Dobricic, S., N. Pinardi, M. Adani, A. Bonazzi, C. Fratianni, and M. Tonani, 2006: “Mediterranean Forecast System: An improved assimilation scheme for sea level anomaly and its validation”, *Quarterly Journal of the Royal Meteorological Society*, in press.

Evans, M., N. Hastings, and B. Peacock, 2000: *Statistical Distributions*, 3rd Ed., John Wiley and Sons, Inc., New York, 221 pp.

PUBLICATIONS

Berliner, L.M., and C.K. Wikle, 2006: “Approximate Importance Sampling Monte Carlo for Data Assimilation”, *Physica D*, in press.

Wikle, C.K., and L.M. Berliner, 2006: “A Bayesian Tutorial for Data Assimilation”, *Physica D*, in press.